

汎用オフィスソフトを利用した アンケートデータの入力と集計⁽¹⁾

内 藤 三 義

〔抄 録〕

アンケート調査データの入力と基本集計について、汎用のオフィスソフト（Excel）を利用して効率的に行うソフトを開発した。これは、データ入力については調査票イメージをそのまま利用したフォームを使用し、マウス操作を基本とする入力を行い、基本集計（属性などによる連続クロス表）についてはピボットテーブル・コマンドを利用し、かつマクロプログラムで整形する方法である。処理に必要な情報は決められた様式でシートに書き込むことで、実現できる。

キーワード アンケート，データ入力，連続クロス表

汎用オフィスソフトを利用したアンケート調査データの入力と集計法について提案する。

筆者はこれまでいくつかの社会調査の実施・集計・分析に関わってきた。その多くが質問紙法（アンケート票）に基づくものであったが、データの入力は、調査ケースの少ない場合などは、調査補助員や自らで入力することもあるが、大部分はデータコーディング法について調整の上、専門エン트리会社に依頼し、集計は統計パッケージソフトを使うことが多かった。しかし、統計パッケージソフトは、高度な分析を進める上では有効だが、分析を深めていく前提として、調査全体の基本集計をコンパクトに表示・印字して、調査メンバーと分かち合う上では不都合があった。

例えば、調査データの初期集計出力として、全ての設問に対する調査対象ケースの基本的な属性ごとの集計出力を表1のような形で出すことになるが、大抵の統計パッケージでは、このようなコンパクトで連続した形での出力はサポートしていない。そこで、今日においてはパソコンにプレインストールされていることも多い汎用オフィスソフトを利用して、こうした出力を得ることができるプログラムを検討してみた。なお、この過程で、データ入力についても、同一アプリケーション上で、基本としてマウス操作でもできるような方法も検討してみた。以下、筆者の開発・提案するプログラムに沿って、データのコーディングから、データ入力、そして基本（クロス表）集計の出力法までについて説明する。

表 1 Excel 集計表例

問 1. 健康状態

	1. 健康	2. 少し悪い	3. 寝ては いない	4. 寝たり 起きたり	5. 寝たきり	無回答	ケース数
全 体	21.3	24.1	46.8	6.7	0.7	0.4	282
地区別							
ひよどり台	13.6	19.3	58.0	8.0	1.1	0.0	88
HAT 灘の浜	24.7	26.3	41.8	6.2	0.5	0.5	194
性別							
男	22.4	25.9	44.8	4.9	1.4	0.7	143
女	20.3	22.5	48.6	8.7	0.0	0.0	138
年齢別							
50 歳未満	46.7	40.0	13.3	0.0	0.0	0.0	15
50-59 歳	40.0	40.0	20.0	0.0	0.0	0.0	15
60-64 歳	29.0	29.0	35.5	3.2	0.0	3.2	31
65-69 歳	20.3	20.3	47.5	10.2	1.7	0.0	59
70-79 歳	14.8	24.3	54.8	5.2	0.9	0.0	115
80 歳以上	19.1	14.9	53.2	12.8	0.0	0.0	47

問 2. 通院, 往診を受けているか

	1. かかって いない	2. 通院し ている	3. 往診	4. その他	ケース数
全 体	21.6	69.9	4.3	4.3	282
地区別					
ひよどり台	15.9	78.4	3.4	2.3	88
HAT 灘の浜	24.2	66.0	4.6	5.2	194
性別					
男	27.3	67.1	0.7	4.9	143
女	15.9	72.5	8.0	3.6	138
年齢別					
50 歳未満	60.0	33.3	0.0	6.7	15
50-59 歳	40.0	46.7	0.0	13.3	15
60-64 歳	29.0	61.3	0.0	9.7	31
65-69 歳	23.7	66.1	5.1	5.1	59
70-79 歳	11.3	80.9	5.2	2.6	115
80 歳以上	21.3	72.3	6.4	0.0	47

「阪神・淡路大震災復興公営住宅入居者, 生活実態調査」⁽²⁾より

1. データの種類と入力

1-1. 回答データの種類とデータ入力のためのコーディング

アンケート票などで回答を求め, 得られるデータとしては, ①選択肢として指定された数値や文字番号, ②実数値, ③自由記入などの文字列がある。また, データ項目（設問・変数）の性質としては, (a) サブ・クエスチョン (Sub Question) と (b) 多重 (複数) 回答がある。データの入力, そのコーディング法を考えるときはこの①~③と (a) (b) の関係を考慮して行うことになるが, 一般的には, 1 ケース 1 レコード (表計算ソフトでは 1 ケース 1 行であ

り、以下これを前提として説明)として、次のようにコーディングし入力することとなる。

①選択肢番号での回答：多重回答でなければ、1設問の回答を1つの列に対応させ、その列のセルに入れる値は選択肢番号である。なお、選択肢番号は文字の場合もあるが、数値に置き換えるほうが都合がよいので、ここでは全て1から始まる整数値として考えている。

②実数値の回答：多重回答でなければ、一般に1設問を1列に対応させ、その列のセルには実数値をそのまま入力する。

③文字列での回答：多重回答でなければ、一般に1設問を1列に対応させ、その列のセルに入れる値は書き込まれた文字列となる。

(a) サブ・クエスチョン：多重回答でなければ1設問を1列に対応させることになる。データ(変数値)の性質により①～③に対応する値をその列のセルに入力する。

(b) 多重回答：1設問に対し複数の列を設定しデータ(変数値)を入力することになる。①の選択肢番号での回答の場合、複数選択方式(選択肢に対して2つまで回答せよなどの設問)では、その数に対応した列数を入力エリアとして確保し、入力するのは選択肢番号となる。複数択一(2値)方式(選択肢に対し幾つでもと回答を求める方式で、各選択肢に対して「該当する」「該当しない」の2値で回答がなされているとみなすことができる)の場合、選択肢番号の数だけ入力エリアが必要となる(全ての選択肢に非該当であるケースと無回答を区別するためにはさらに無回答や非該当を一つの選択肢として設定することも必要である)。なお、複数選択方式であっても、データの入力やコーディングは複数択一(2値)方式として扱っても問題はない。また、その方が特に非該当と無回答が混在するデータの場合、集計上は都合のよいことが多い⁽³⁾。

以上を考慮し、1行を1ケースに対応させることとしながら、どの列にどの設問の回答を入力するのか、またその入力データがどのような性質のデータであるのかを吟味して入力レイアウトとフォーマットを決めることになる。

アンケート票例を図1に、そのデータ入力フォーマットを表2に示している。最初の問いは選択肢番号を1つだけ回答する設問なので1セルに数値を入力、問2も同様だが、「その他」に記入がある場合はそれを1セルに文字列として入力する。問2-1、は複数回答なのでここでは2値方式で、選択肢の数だけセルをあてはめ、該当する選択肢が選ばれているとき、そこに「1」という値を入力する方式をとっている。多重回答設問に無回答の場合は、それも一つの選択肢として扱うため1セルとっている(これはデータとして入力せずに、データ入力後に論理演算などで値を当てはめることもできる)。冒頭にアンケート票にはない、IDの項目の入力をデータ点検・識別用に1セルとっているので、この見開き2頁のデータを入力するには68セルが必要ということになる(図1と表2では多重回答をすべて2値方式として入力しているが、問9、問10などを3つまでの複数選択方式として入力すれば、43セルとなる)。

図1 アンケート票の例

ID:

最初に、あなたの健康状態のことなどについてお聞きします。

問1. あなたの健康状態はどうか。(○を一つ)

☐ 1. 健康 (特に悪い所はない) ☐ 2. 少し調子が悪い
☐ 3. 寝てはいいが朝気を持っていない ☐ 4. 寝たきり近き状態
☐ 5. 6ヵ月以上床につききり

問2. 現在、あなたは病院や診療所などに通院、または往診を受けていますか。(○を一つ)

☐ 1. 医療機関にかかっていない ☐ 2. 通院している
☐ 3. 往診してもらっている ☐ 4. その他

問2-1. 通院・往診を受けている方にお聞きします。利用している医療機関はつぎのうちどれですか。(あてはまるものすべてに○)

☐ 1. 今の住居近くの病院・診療所
☐ 2. 家族など親類住宅入居以前に住んでいたときの病院・診療所
☐ 3. 震災前からのかかりつけの病院・診療所
☐ 4. その他

問3. あなたの世帯の1ヶ月の医療費(通院のための交通費なども含む)は、おおよそいくらでしょうか。金額をお書きください。

円

問4. 利用している医療保険の種類は何ですか。(○を一つ)

☐ 1. 国民健康保険の本人 ☐ 2. 国民健康保険の家族
☐ 3. 社会保険の本人 ☐ 4. 社会保険の家族
☐ 5. 生活保護の医療扶助 ☐ 6. 医療保険に入っていない
☐ 7. その他

問5. 医療費の負担は大変ですか。(○を一つ)

☐ 1. 大変重いと感じている ☐ 2. 少し負担だと感じている
☐ 3. 負担だとは思わない ☐ 4. その他

問6. あなたや、同居している方の中に、介護の必要な方はおられますか。(あてはまるかたすべてに○をつけてください)

☐ 1. 自分 ☐ 2. 配偶者 ☐ 3. 親 ☐ 4. 兄弟
☐ 5. 子供 ☐ 6. その他の同居者 ☐ 7. 介護が必要なものはない

問7. 来年4月から介護保険制度が始まりますが、あなたの世帯では、その申請をしようと思っていますか。(○を一つ)

☐ 1. すでに申請した ☐ 2. 申請していないが、申請したい
☐ 3. 負担が重いので、申請しない ☐ 4. 必要ないので申請しない
☐ 5. その他

問8. 来年4月からの介護保険は、65歳以上の人の場合、神戸市では平均月3270円になると予定されています。この金額についてあなたはどうか考えますか。(○を一つ)

☐ 1. 高すぎる ☐ 2. ちょうどよい
☐ 3. 安い ☐ 4. わからない
☐ 5. その他

次に、あなたの暮らしの状況についてお聞きします

問9. あなたの世帯の収入源は何ですか。(主なもの3つまで○)

☐ 1. 年金などの貯蓄収入 ☐ 2. 売上の営業収入 ☐ 3. 公的年金
☐ 4. 企業年金など私的年金 ☐ 5. 預貯金の引き出し ☐ 6. 賃貸・利子・配当
☐ 7. 子どもなどの仕送り ☐ 8. 同居家族の収入 ☐ 9. 生活保護
☐ 10. その他

問10. あなたの世帯で受けている公的年金の種類は何ですか。(ご家族の人も含め、あてはまるものすべてに○)

☐ 1. 厚生年金 ☐ 2. 共済年金 ☐ 3. 国民年金 ☐ 4. 恩給
☐ 5. 遺族年金 ☐ 6. その他 7. 受けていない

問11. 世帯の収入総額はどれくらいですか。月額でお書きください。(○を一つ)

☐ 1. 5万円未満 ☐ 2. 5~8万円未満 ☐ 3. 8~10万円未満
☐ 4. 10~13万円未満 ☐ 5. 13~15万円未満 ☐ 6. 15~20万円未満
☐ 7. 20~25万円未満 ☐ 8. 25~30万円未満 ☐ 9. 30万円台
☐ 10. 40万円以上

問12. 毎月の生活費でかさんで困っているものに、3つまで○をしてください。

☐ 1. 主食費 ☐ 2. 副食費 ☐ 3. 娯楽以外の教科 ☐ 4. 酒類
☐ 5. 外食費 ☐ 6. 家具・家庭用品 ☐ 7. 衣料費 ☐ 8. 水道代
☐ 9. 電気代 ☐ 10. ガス代 ☐ 11. 保険・医療費 ☐ 12. 交通費
☐ 13. 通信費 ☐ 14. 娯楽費 ☐ 15. 献費・献費 ☐ 16. 交際費
☐ 17. ローンなど借金返済 ☐ 18. その他
☐ 19. 特に出ているものはない

- 2 -

途中終了

- 3 -

次の頁へ

前掲、「阪神・淡路大震災復興公営住宅入居者、生活実態調査」より。なお、この図は Excel のフォームに変換したもので、既にオプションボタンやテキストボックスなどが入っている。ボタンなどが無いのが、元々のアンケート票である。

表2 図1のアンケート票に対応した入力フォーマット例

ID	問 1	問 2	問 2-1 (S. M)						問3 (R)	問 4	問 5	問 6 (M)								問 7			
			記入	1	2	3	4	NA	記入				記入	1	2	3	4	5	6	7	NA		記入
N	N	N	S	N1	N1	N1	N1	N1	S	N	N	N	S	N1	N1	N1	N1	N1	N1	N1	N1	N	S

問 8	問 9 (M)												問 10 (M)								
記入		1	2	3	4	5	6	7	8	9	10	NA	記入	1	2	3	4	5	6	7	NA
N	S	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	S	N1	N1	N1	N1	N1	N1	N1	N1

問11	問 12 (M)																				
1		2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	NA	記入
N	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	N1	S

N は数値を入力するセル, S は文字列, N1 は該当する場合「1」を入力するセル。

1-2. 非該当と無回答など欠損値の処理方法

アンケート方式の調査では、回答者が特定の設問に対して回答をしないこともある。そうした「無回答」による入力データがない場合と、サブ・クエスチョンなどで回答が不要なケースにあたる場合の「非該当」は集計・分析上峻別する必要がある。データ入力の際は、これを

初めからコーディングして入力するか、入力後に論理的に判断してデータの整形を行うか、決めておく必要がある。なお、このことについては、調査票設計時にも注意がいる。選択肢番号で回答を求める設問の場合、全ての選択肢が回答者のケースに該当しないために回答をしないケースが「無回答」とならないようにすることである。「どれもあてはまらない」「その他」などの選択肢が必要な設問があることに留意する必要がある。最終的なコーディングとしては、筆者は通常、選択肢式の設問の非該当と、実数値で回答を求める設問の無回答は半角ピリオドとし、数値による集計の対象としないようにしている。また、選択肢式の設問でクロス集計などに反映する無回答は99や999、などに設定している。

1-3. データの入力

汎用計算機までの時代に戻れば、アンケートデータのファイル形式と入力形式は、固定長・固定書式が一般的であった。しかし今日では汎用オフィスソフトのファイル形式で入力し、必要な場合にCSV (Comma Separated Values) やテキスト形式に変換するのが普通のこととなっている。データ入力会社に依頼する場合、それぞれの会社独自の入力支援プログラムが利用されることも多いと思われるが、データの受渡しはExcel形式やCSV形式となることが多くなっている。ところで、Excelなどの表計算ソフトを利用したデータ入力としては、該当する行・列のセルに移動してキーボードで数値や文字列を入力するのが普通であるが、アンケートデータの入力のように、1ケース(1行)のデータをセルを移動させながら入力する作業では、特にデータ数(設問数、選択肢数)が多くなると入力ミスが起きやすい。図2を見てもわかるように、アンケート票への記入箇所は平面上の配置で統一がとれていないことが多いので、見落としが起き、そのことが誤ったセル移動での入力を誘う二重のミスを起こしやすいことなどである。そこで、アンケート票の設問や選択肢配置に対応させた入力フォームを用意し、それをデータ入力に利用する入力方法を提案する。

その基本的な考え方は、(1) Excelのフォームを利用する、(2) フォームにはアンケート票の各頁、または見開き頁のイメージを直接利用する(これはフォームにアンケート票と同じようなレイアウトを文字列入力で行うのはそれなりに面倒な作業だからであり、その問題がなければフォームに直接文字入力してもかまわない。イメージは印刷物をスキャンして作成する方法もあるが、アンケート票の版下にあたるものがコンピュータから直接出力できるファイルとなっている場合は、PDF作成ソフトなどを利用したイメージファイル出力をする方が簡単であろう)、(3) 回答にあたる選択肢、文字列入力にはフォームの「オプションボタン」「チェックボックス」「テキストボックス」を利用する、(4) その利用の仕方や数などをコントロールする情報はプログラムではなく、シートの中に記入されたテーブルデータで行う、というものである。なお、選択肢式の設問はすべて1から始まる整数とし、複数回答の2値式の回答のある選択肢には1を、その他はブランクを設定するものとして扱っている。

ほとんどがマウス操作でデータ入力できる形のフォームとして設定したのが図 1 である。選択肢番号からの択一式の回答を求めるものにはオプションボタン（どれか 1 つしか選べないボタン）のグループを設定し、複数回答はチェックボックス（いくつでもチェックできるボックス）のグループを、直接記入する項目にはテキストボックスを割り当てている。選択されたボタン情報やテキストの書き込みは、指定した Excel ファイルの指定シートの該当箇所に、選択肢番号の数値で書き込まれ、テキストも同様に指定セルに書き込まれるようプログラムが組まれている。

なお、図 1 の例は、無回答や非該当を事後に論理的に判断して、整形や変数追加を行うものとしているが、無回答・非該当を 1 つの選択肢として付け加える方法もある。また、択一式の設問でも回答者によっては 2 つ以上を誤って回答することがある場合、どれが正答であるかを入力担当者の判断に任せず、データ入力では誤回答もそのまま反映することとし、択一式の回答もオプションボタンではなく、チェックボックスで対応する（図 1 の問 1 を例にすると、オプションボタンならデータは 1 セルに入力できるが、この方法だと複数回答と同じだけのセルが必要となる）という方法もある。

図 2 データ入力のためのコントロール情報記入例

	1	2	3	4	5	6	7	8	9	10	11	12	13	1
8	フォームの数				4									
9						第1グループ			第2グループ			第3グループ		
10						グループ	データ出力	入力	データ出力	入力	データ出力	入力	データ出力	
						ボタン	開始	チェック	ボタン	開始	チェック	ボタン	開始	ボタン
11	Form1を書出すブック名	DataIn.xls	Form1	OptionButton	7	5	2	0	4	3	0	7	12	0
12	" Sheet名	Data	"	CheckBox	5	4	5	0	7	15	0	10	27	0
13	データ書き出し開始行	6	"	TextBox	9	1	1	1	4	0	1	10	10	0
14	Form2を書出すブック名	DataIn.xls	Form2	OptionButton	9	2	88	0	5	99	0	5	100	0
15	" Sheet名	Data	"	CheckBox	4	18	69	0	7	89	0	12	102	0
16	データ書き出し開始行	6	"	TextBox	4	97	0	1	98	0	1	130	0	0
17	Form3を書出すブック名	DataIn.xls	Form3	OptionButton	13	8	137	0	2	138	0	4	142	0
18	" Sheet名	Data	"	CheckBox	1	15	148	0						
19	データ書き出し開始行	6	"	TextBox	10	135	0	1	136	0	1	139	0	0
20	Form4を書出すブック名	DataIn.xls	Form4	OptionButton	4	2	175	0	3	176	0	5	178	0
21	" Sheet名	Data	"	CheckBox	0									
22	データ書き出し開始行	6	"	TextBox	3	174	0	1	177	0	1	179	0	0

前掲調査（図 1）に対応したコントロール情報の記入例。このアンケート票は本文が 7 頁（見開きを 1 頁とすると 4 頁）であり、入力フォーム数は 4 つとなっている。第 1 フォームでは、オプションボタンを割り当てた択一式の設問が 7 つ、問 1 は選択肢が 5 つで、そのデータは第 2 列に書き出される。多重回答（2 値方式）のためチェックボックスを割り当てた設問数は 5 つで、その最初の問 2-1 は、選択肢が 4 つで、データの書き出しは 5 列以降。直接入力のためテキストボックスを割り当てたのは 9 つで、最初の ID は 1 列に書出すが、ID は必ず書き込まなければならないので、入力チェックを 1 とし、書き込みが無ければ警告を出すようにしている。

フォームの入力情報を読み込み、データシートに書出すプログラムは、アンケート毎に作り直す必要はないようにしている。基本的な情報は全て Excel のシートに図 2 のように記入することでよく、アンケートの入力設定と集計を行うユーザが設定するのは、(1) 表 2 に示したようなデータのコーディングと入力フォーマットを作成する、(2) アンケート票のイメージ

出力などを準備した **Excel** のフォームに、図 1 のように各設問の選択肢に対応するオプションボタン等を指定規則どおりに配置する、(3) 入力フォーマットとアンケート票に対応した図 2 のようなコントロールテーブルをシートに作成する、(4) フォームが複数にわたる場合は、フォームに対応する **Excel** マクロをコピーし、プログラムの冒頭に指定されているフォーム番号の数字を修正するという作業である。

データ入力を実際に行う担当者は、こうした作業とプログラムがどのようになっているか特に認識する必要はなく、**Excel** を起動してこのマクロとフォームのあるファイルを読み込んだ後、アンケート表に対応した図 1 に示した入力シートの該当箇所をマウス等でクリックすればデータ入力が行進する。

2. 入力データの整理

上記のような方法で入力されたアンケートデータは、そのまま直接集計に利用できる場合もあるが、様々な整形・変換をしなければならないことが多い。特にデータ入力にかかる労力を削減しようとして、無回答や非該当などはデータを入力せずに、入力データから論理的に判断しようとする場合、そうした判断によるデータの追加・整理などが必要となる。また、実数値で回答を求めて、それをカテゴライズしたデータが必要となる場合（例えば年齢を実数で回答を求め、5 歳刻みの年齢階層の集計をする場合）なども、そうして作成したデータの追加が必要となる。

3. 連続クロス表の作成

上記の作業が終了することで、集計作業に入り、表 1 で示したような連続クロス表の作成に移るが、**Excel** の一般的な操作で、直接このような集計ができるわけではない。**Excel** にはピボットテーブル・レポートという集計表作成の機能があるが、これはクロス表を作成できるが、連続した表を 1 回の操作で作ることはできない。また、ケース数の表は作成できるが、% 値の表はこのコマンドだけでは作成できない。そこでピボットテーブルを連続して作成し、必要に応じて % 値の計算、設問番号や選択肢番号のラベルへの置き換えなどを実行できる **Excel** のマクロプログラムを作成した。

このプログラムがめざしたのは、データ入力プログラムと同様、読み込むデータの情報や、出力したい変数の指定などを図 3 のように **Excel** のシートに書き込めば、それを参照しながらクロス表を作成していくようにする、ということであり、ユーザが作成するのは、①入力されたデータ (**Excel** 形式)、②そのデータファイルの中の集計対象となるデータの範囲や出力したいクロス表の変数群の入力列を指定した **Excel** ファイルのシートである。

図 3 集計のための情報シートの例

1	2	3	4	5	6	7	8	9	10
17									
18	a クロス表作成							c パーセン	
19									
20	データファイルのブック名		Sample.xls					実数値から%値のクロス表に変換	
21	その中のデータのシート名		Data					データファイルのブック名	
22								" の中のシート名	
23	テーブルの開始・行番号		4	相対列番号とは、図格列から見て何列目が					
24	" 終了・行番号		104	ということ(左が「1」なら、絶対番号である)					
25	テーブルの開始・列番号		1						
26	" 終了・列番号		164	ダミー変数の相対列番号 164					
27								m 平均	
28									
29	基準変数について Max=100			分析変数について Max=100				平均値の作成	
30									
31	その数		4	その数		39		その数	3
32				多重回答の場合その変数の数					
33									
34	第1基準変数の相対列番号		163	第1変数の相対列番号		2		第1変数の相対列番号	9
35	第2 "		160	第2 "		3		第2 "	82
36	第3 "		161	第3 "		4	5	第3 "	151
37	第4 "		162	第4 "		157		第4 "	
38	第5 "			第5 "		10		第5 "	
39	第6 "			第6 "		11		第6 "	
40	第7 "			第7 "		12	8	第7 "	
41	第8 "			第8 "		20		第8 "	
42	第9 "			第9 "		21		第9 "	
43	第10 "			第10 "		22	1003	第10 "	
44	第11 "			第11 "		25	8	第11 "	
45	第12 "			第12 "		33		第12 "	
46	第13 "			第13 "		34	20	第13 "	

前掲調査（図 1）に対応した出力情報の記入例でサンプルデータに即している。このデータでは、「その他」の書き込みなどは割愛しており、また問 9 の多重回答は複数選択方式（MA 型）で入力されている。データは変数名の書き込まれた 4 行から 104 行までで、1 列から 164 列までが集計対象であると指定している。作成するクロス表は、163（総数）、160（地区）、161（性）、162（年齢）の変数の列と 2 列（問 1）のクロス表、次いで前者と 3 列（問 2）のクロス表であるというような指定になっている。

プログラムが実行しているのは、図 3 のような形で指定された情報で、(1) ピボットテーブルの範囲を設定し、(2) 出力した変数間のクロス表をピボットテーブル・コマンドで 1 つ作成する（最初は第 1 基準変数と第 1 分析変数のクロス表）、(3) それを出力用シートにコピーする形で書き出したあと 第 2 基準変数と第 1 分析変数のクロス表の作成と、それを出力シートの先ほど書き出した後ろの行にコピーするという形で (2) と (3) を繰り返す、というものである。ここで得られた出力は基本的にはピボットテーブル・コマンドの「データの個数」による集計なので、欠損値の処理、%値の計算、変数や変数値に対する（日本語）ラベル付けなどはしていない。そこで、それらの処理についても実行するプログラムを準備したが、それらも、処理に必要な情報は Excel ファイルのシートに書き込むと、それを参照して処理が進むようにしている。なお、(3) の処理の際、連続クロス表の各行の意味を表すインデックスを付けた列を 2 列確保し、追加処理の際に利用している。

基本となるクロス表の作成で特に工夫が必要となったのが、多重回答の処理である。これに

については2値式(YN型)の場合は、各変数(列)について「1」の値のみを集計し、それを繋ぐという処理となった。また、複数回答式(MA型)の場合は準備した列について順に集計しながら、それぞれの列での選択肢への回答数を、変数値の値ごとに足し合わせ、昇順ソートをかけるというプログラミングで実現した。

なお、このプログラムには連続平均値表も作成できるようにしているが、その考え方と、処理の流れはクロス表と同様である。

お わ り に

連続クロス集計プログラムは既に2004年にはネット上に公開している。今回、データ入力フォーム用のプログラムも開発する中で、クロス出力プログラムも一部を見直し改訂することで、この研究ノートを発表することとなった。

アンケート集計の初期作業を出来るだけ効率的に実行したいというのが、今回の提案の基本にある。データ入力のためのフォームを利用したプログラムは、データエントリーを調査補助員など、「自前」で行う場合のことを考えて作成した。調査ケース数が100ケースに満たない場合などは、フォームを作成する手間を考えると、あまり省力化にはならないであろうが、入力ケース数が数百になり、分担してデータ入力を行う時などは非常に有効なデータ入力ソフトになるであろう。

連続クロス表作成プログラムは、ケース数に関わりなく効率的な初期集計作業に十分活用できるものと確信する。このプログラムでの集計で、準備作業に最も時間を取られるのはクロス表の変数や変数値にラベル付けをするための文字入力作業であり、それが不必要ならそれほどの手間無く集計が得られる。またラベル付けのためのシートへの入力も、アンケート票の原稿ファイルを利用して、切り貼りを効率よく行えば作業時間も短縮できる。

この2つのプログラムの提案と提供が、社会調査などに携わる者にとって、初期集計作業を効率的に進めることで、分析の質を高めることにつながればと期待するものである。

なお、このプログラムはExcel利用を前提にしているが、OpenOffice.orgのCalcなど、Excel以外の無償で入手できるソフトへの対応も考慮に入れているが、次の開発作業に任せたい。

〔注〕

- (1) 本稿で説明しているデータ入力や連続クロス表作成プログラムと、その利用の仕方の詳細は筆者のホームページ(<http://www.bukkyo-u.ac.jp/mmc/01/naito/>)で既に公表している。本稿はそのエッセンスを要約的に示したものである。

なお、本稿で使用したExcelはMicrosoft Corporationの登録商標である。

- (2) このアンケート票は、阪神・淡路大震災復興公営住宅入居者生活実態調査実行委員会が1999年

11月20・21日に実施した「阪神・淡路大震災復興公営住宅入居者，生活実態調査」に使用したものである。少し古いデータであるが，この調査データから，集計用のサンプルのデータ100ケース分，連続クロス表作成のためのコントロールシートなどを筆者ホームページで提供しており，一貫性を考慮しこのデータに即して説明することとした。

- (3) 多重回答データ整形法などについては，拙稿「多重回答データの編成と統計処理」，立命館産業社会論集，39号，1984年，奥川桜豊彦と共著，ですでに説明している。なお，この論文では複数選択式は「Multi Answer (MA) 型」，2値方式は「Yes or No (YN) 型」と名づけて説明している。

〔付記〕

この研究ノートは，2006年度佛教大学国内研修の成果として発表している。

（ないとう みつよし 現代社会学科）
2008年10月2日受理